Nature Conservation 2: 41–57 (2012)
doi: 10.3897/natureconservation.2.3644
http://www.pensoft.net/natureconservation

RESEARCH ARTICLE

A peer-reviewed open-access journal
Nature Conservation
Launched to accelerate biodiversity conservation

Bird-monitoring in Europe – a first overview of practices, motivations and aims

Dirk S. Schmeller[1,2], Klaus Henle[2], Adeline Loyau[1,2], Aurélien Besnard[3], Pierre-Yves Henry[4]

1 *Station d'Ecologie Expérimentale du CNRS à Moulis, 09200 Saint Girons, France* **2** *UFZ - Helmholtz Centre for Environmental Research, Department of Conservation Biology, Permoserstr. 15, 04318 Leipzig, Germany* **3** *Centre d'Ecologie Evolutive et Fonctionnelle, UMR 5175, 1919 Route de Mende, F34293 Montpellier Cedex 5, France* **4** *UMR 7204 & UMR 7179 CNRS MNHN UPMC, Centre de Recherches par le Baguage des Populations d'Oiseaux, Muséum National d'Histoire Naturelle, 55 rue Buffon, 75005 Paris, France*

Corresponding author: *Dirk S. Schmeller* (ds@die-schmellers.de)

Academic editor: *Dmitry Geltman* | Received 8 July 2012 | Accepted 23 August 2012 | Published 24 August 2012

Citation: Schmeller DS, Henle K, Loyau A, Besnard A, Henry P-Y (2012) Bird-monitoring in Europe – a first overview of practices, motivations and aims. Nature Conservation 2: 41–57. doi: 10.3897/natureconservation.2.3644

Abstract

Biodiversity monitoring is central to conservation biology, allowing the evaluation of the conservation status of species or the assessment of mechanisms of biodiversity change. Birds are the first taxonomic group to be used to build headline indicators of biodiversity due to their worldwide spatial and temporal coverage and their popularity. However, the landscape of bird-monitoring practices has never been characterized quantitatively. To objectively explore the strengths and weaknesses of the massive bird-monitoring effort in Europe we assessed the bird-monitoring practices, acquired with a questionnaire-based survey, in a sample of monitoring programs. We identify major correlates of among-program variability and compare monitoring practices from our database to recommendations of best monitoring practices. In total, we obtained responses from 144 bird-monitoring programs. We distinguish three types of monitoring programs according to the number of people that they involve: small, local-scale programs (56%), medium or regional programs (19%), and large-scale, national and international, programs (23%). In total, the programs in our sample involved 27941 persons, investing 79298 person days per year. Our survey illustrated that 65% of programs collected quantitative indices of abundance (count data). The monitoring design in a majority of the programs could be improved, notably in terms of unbiased spatial coverage, sampling effort optimization, replicated sampling to account for variations in detection probability, and more efficient statistical use of the data. We discuss the main avenues for improvement in bird-monitoring practices that emerge from this comparison of current practices and published methodological recommendations.

Copyright *Dirk S. Schmeller et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC-BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Keywords
conservation priorities, biodiversity monitoring, volunteers, survey design, conservation status in Europe, conservation policy support

Introduction

Biodiversity and environmental monitoring provide fundamental information for tracking environmental changes, to diagnose population trajectories and to provide conservation biology with relevant data. Such information is required for the design and evaluation of biodiversity policies, conservation management, land use decisions, and environmental protection. Biodiversity monitoring is therefore central to conservation biology, allowing the evaluation of the conservation status of species or to assess biological responses to environmental changes (such as climate change, Lepetz et al. 2009), and to conservation policy (Male and Bean 2005; Taylor et al. 2005; Donald et al. 2007).

A large number of monitoring programs have been developed and a large body of literature on biodiversity monitoring is available, including several articles that provide recommendations for an optimal design of monitoring programs (Danielsen et al. 2000; Yoccoz et al. 2001; Kery and Schmid 2004; Vořišek et al. 2008; Lindenmayer and Likens 2009; 2010). Apart from methodological advice, most of these articles agree that many monitoring programs were poorly designed and, therefore, could be a waste of time and resources (Nichols and Williams 2006). However, quantitative assessments of monitoring practices at varying spatial scales were not available at the time of these publications (Marsh and Trenham 2008). Large databases collecting data on and rating monitoring practices are now becoming available (Kull et al. 2008; Lengyel et al. 2008; Schmeller et al. 2009) and provide the first opportunity for a quantitative assessment of how well monitoring practices match methodological recommendations.

Bird-monitoring initiatives are the first provider of long-term monitoring data when institutional bodies set the goals of quantifying global biodiversity changes and of assessing the impact of environmental policies on biodiversity (Tucker and Heath 1994; Burfield et al. 2004; Gregory et al. 2005; 2006). In many instances, birds are the taxonomic group for which most data are available. Hence, we should characterize monitoring practices, and develop recommendations of how they could be improved for an optimized future monitoring effort. Further, such an assessment of the state of biodiversity-monitoring practices may contribute to the establishment of a global monitoring system, as envisaged by the Group of Earth Observation – Biodiversity Observation Network (GEOBON; Pereira et al. 2010).

For the first time, a comprehensive database of the FP6-project EUMON (hereafter DaEuMon; Schmeller et al. 2009) made available standard information describing biodiversity-monitoring practices in Europe. This meta-database contains data on sampling practices, sampling efforts, sampling design, volunteer involvement etc., of

600 European monitoring programs and aims at describing the monitoring landscape in Europe. Here we used this data source, to characterize bird-monitoring practices. We focused on differences among programs in motivation and aims, sampling design, sampling effort and methods of data analysis during the monitoring process. Further, we analyzed differences in these parameters among bird species groups (raptors, songbirds and near passerines, waterbirds), and according to the size of a monitoring program as defined by the number of persons involved. Our characterization of the overall European landscape of bird-monitoring practices will address the general questions: What are the average practices of bird-monitoring in Europe? And how do these practices relate to the motivation and aim for monitoring, to sampling effort, and to the involvement of non-professionals? A summary of this information will act as an aid for those wanting to launch a new program, improve the design of an ongoing monitoring program (adaptive monitoring), or evaluate bird-monitoring data quality. Our approach differs from earlier publications focusing only on national or international federations of monitoring programs (Gibbons 2000; Vořišek and Marchant 2003; Klvanová and Voříšek 2007) as we also include regional and local monitoring programs.

Methods

600 monitoring programs are available in DaEuMon database; They were obtained through a questionnaire survey (ESM1). Among them, 144 concern bird species and were analyzed in detail. We checked responses for completeness, and sought missing details from the coordinators of monitoring projects. Once the responses have been validated, data were made publicly available through our online database (http://eumon.ckff.si/biomat/). Complete information was not available for every single question for all programs, hence affecting sample sizes in the analyses.

For the characterization of the bird-monitoring landscape, we focused on differences in the motivation and aims, sampling design, sampling effort and methods used for data analysis. We analyzed the differences between bird species groups (raptors, songbirds and near passerines, waterbirds) and between monitoring programs of different sizes in terms of the number of persons involved. We defined three size-categories: small ($N_{persons} \leq 30$; N = 81), medium ($N_{persons}$ 31 – 150; N = 26) and large monitoring programs ($N_{persons} > 150$; N = 32). The motivation was characterized by the program objective (scientific, management or political/juridical), the type of trends monitored (distribution, population size or avian community trends) and the focal ecological factor (climate change, habitat fragmentation, pollution, invasive species, land use). Sampling design was characterized by site choice methodology, the use of stratified sampling or not, the use of repeated sampling or not (that allow accounting for detection probability), the location of sampling sites within and/or outside protected areas, and the main field data type collected (Presence/absence, Counts, Mark recapture, Age/size structure, Phenology). We further quantified the sampling effort by the number of

species ($N_{species}$), persons ($N_{persons}$), sites (N_{sites}), visits per site (N_{visits}), sampling effort in person.days, and the proportion of volunteers (%Vol).

We tested for differences in practices with generalized linear models (GLM) using SAS 9.1.3 (Cary, USA, 2002); GLM with a multinomial distribution of error terms and a clogit link function for the type of field data (categorical variable); GLM with a Poisson distribution of error terms and a log link function for the number of species monitored; GLM with a binomial distribution of error terms and a logit link function for the analysis of the use of stratification, of detection probability, and of advanced statistics. The dependent variables were therefore: the type of field data, the number of species monitored, the use of stratification, the use of detection probability, and the use of advanced statistics. The corresponding independent variables included in the models were: the number of persons involved in the program, the number of professionals, the ratio of volunteers, the number of person days, and the program objective. We also included the sampling design used when analyzing the use of advanced statistics. The models were adjusted for overdispersion when necessary. We conducted a stepwise procedure with a backward elimination at the 5%-level, starting with a fully saturated model, incorporating all independent variables with no interaction, and dropping, step-by-step, all non-significant variables. At each step, the term that gave the smallest contribution to the model (largest p-value) was excluded.

Bias in geographic coverage

A major problem of surveys such as ours (volunteer response to a mailed questionnaire) is that it is nearly impossible to achieve a random sample because of the decentralized structure of the network of monitoring activities (Schmeller et al. 2009 for Europe and Marsh and Trenham 2008 for North America). Indeed, monitoring coordinators of highly visible monitoring programs have a certain fatigue toward questionnaires or strong time constraints and simply may not reply (Barclay et al. 2002). At the opposite end of the size gradient, it is hard to get in touch with a large number of local, non-federated monitoring programs, which represent a large subset of the available monitoring data. The EuMon survey encountered both problems. For example, not all coordinators of national Breeding Bird Surveys (listed on the page of the Pan-European Common Bird-monitoring Program, PEBCM; EBCC 2010) contributed to the EuMon survey. A direct comparison of programs covered by both surveys is difficult, as the EuMon survey covered a much larger range of different monitoring programs (from local to international) than the EBCC list, which focuses on national programs only. Further, titles of national programs differed between the EBCC and EuMon surveys.

Despite a large effort in sending out requests for cooperation to a wide audience, our survey data provide a characterization of monitoring practices in Europe that suffers from a biased geographic coverage. We used GoogleScholar to estimate the bias

in our sample by looking for articles with the search string ("bird-monitoring" OR "bird survey" country). Our analysis shows that Lithuania, Poland, France, Bulgaria and Andorra were overrepresented in our program, while Great Britain, for example, was underrepresented (Figure 1). Also in comparison to data collected by EBCC, our survey has obviously undersampled bird-monitoring programs in Great Britain and Sweden. Our survey covers 24 European countries, with a strong (over-) representation of France and Poland (Figure 1; Schmeller et al. 2009). Despite this non-random coverage of European countries, our database is for now the most extensive data set to characterize bird-monitoring practices in Europe. Other initiatives analyzing bird monitoring programs focused on large-scale, national breeding bird surveys (Gibbons 2000; Voříšek and Marchant 2003; Klvanová and Voríšek 2007), which may be considered as the most visible and legitimate minority within the whole bird-monitoring community.

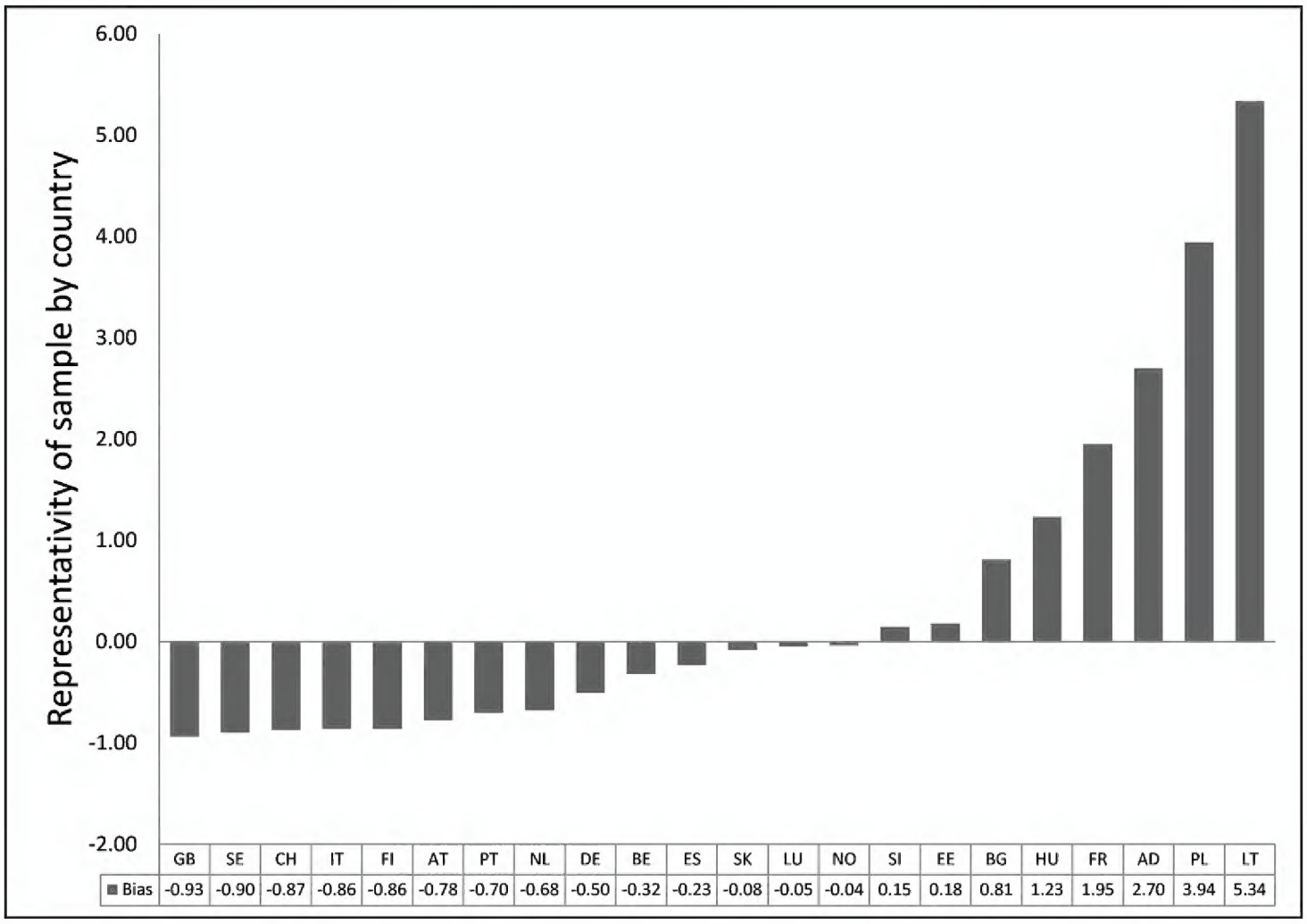

Figure 1. Estimation of the bias in the number of bird-monitoring programs in the EuMon database per country (bias = [Number of programs DaEuMon – Number of articles in Google Scholar]/ Number of articles in Google Scholar). The reference to quantify bird monitoring activity per country was the number of publications in GoogleScholar returned for the search string ("bird-monitoring" OR "bird survey" AND country name). The countries are abbreviated following the two-letter convention of the international community (ISO 3166-1 alpha-2 codes; GB = Great Britain, SE = Sweden, CH = Switzerland, IT = Italy, FI = Finland, AT = Austria, PT = Portugal, NL = Netherlands, DE = Germany, BE = Belgium, ES = Spain, SK = Slovakia, LU = Luxembourg, NO = Norway, SI = Slovenia, EE = Estonia, BG = Bulgaria, HU = Hungary, FR = France, AD = Andorra, PL = Poland, LT = Lithuania).

Results

Our European-wide survey yielded responses from 144 bird-monitoring programs employing 27941 persons investing 79298 person days per year. The majority of responses recorded in our database came from France (49; 34%) Poland (28; 19%), and Lithuania (13; 9%). Six to eight responses came from the Netherlands, Germany, Spain, Norway, and Hungary (35; 24%; Figure 2). In total, all bird-monitoring programs employed 27941 persons, with a mean of 201+/- 75 persons per program and a mean manpower of 615 +/- 138 person days per year per program.

Small programs monitored 29 +/- 5.2 species (median = 6) with 11 +/- 1 persons (median = 7), which were mainly professionals (66% +/- 4.5), investing on average 148 +/- 34.2 days per year, visiting 116 +/- 67.7 sites (median = 12) on average 9.4+/-2.9 times (median = 3). Medium sized programs monitored 42 +/- 17.8 species (median = 5) with 64 +/- 4.7 persons (median = 60), which most frequently were volunteers (77% +/- 5.7), investing on average 492 +/- 109 days per year, visiting 67 +/- 16.9 sites (median = 43) on average 28.7+/-19.1 times (median = 9.5). Large programs monitored 72 +/- 12.2 species (median = 70.5) with 793 +/- 306 persons, of which most

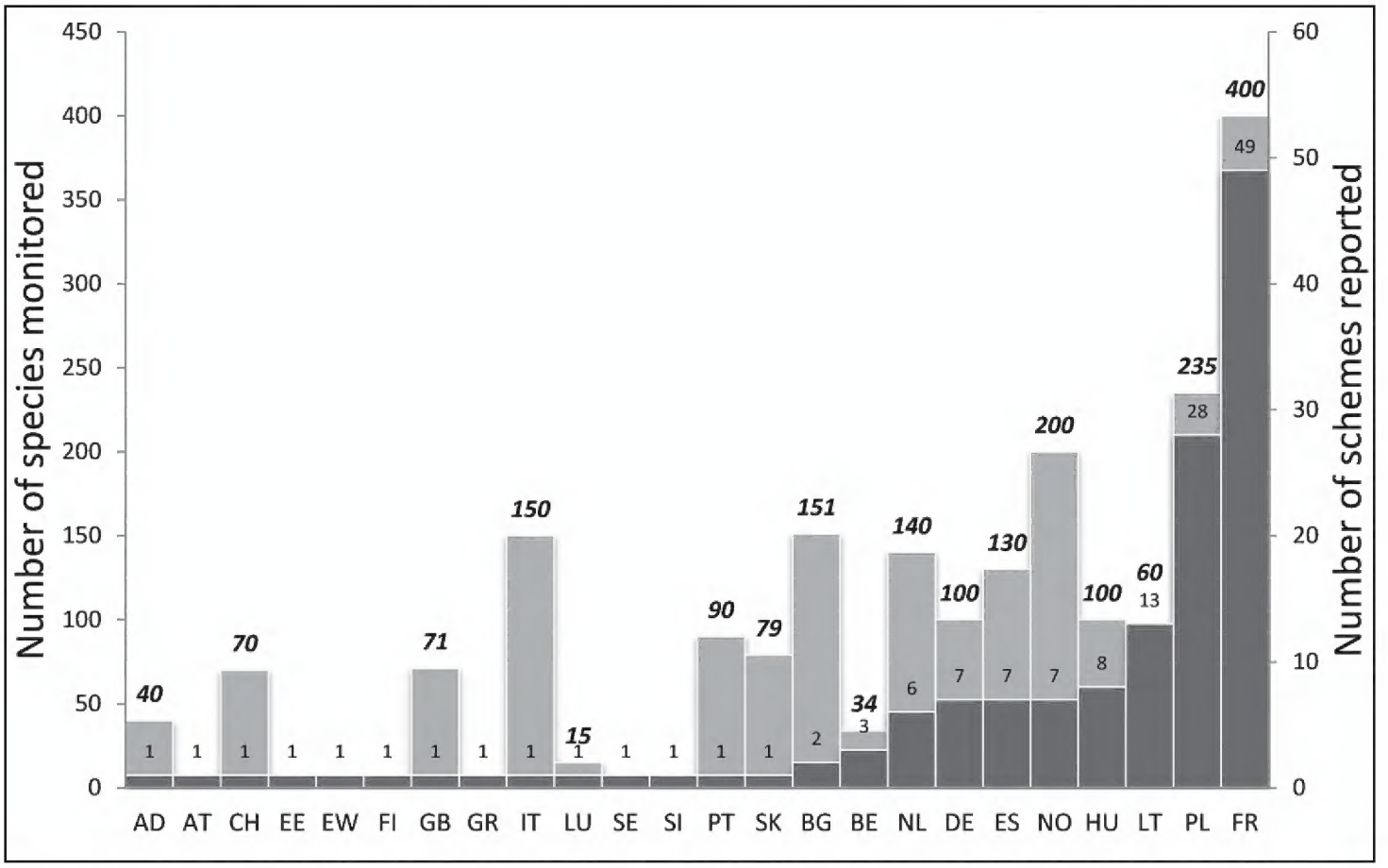

Figure 2. Number of responses (dark grey) and number of species monitored (light grey) per country. The countries are abbreviated following the two-letter convention of the international community (ISO 3166-1 alpha-2 codes). (AD = Andorra, AT = Austria, CH = Switzerland, EE/EW = Estonia, FI = Finland, GB = Great Britain, GR = Greece, IT = Italy, LU = Luxembourg, SE = Sweden, SI = Slovenia, PT = Portugal, SK = Slovakia, BG = Bulgaria, BE = Belgium, NL = Netherlands, DE = Germany, ES = Spain, NO = Norway, HU = Hungary, LT = Lithuania, PL = Poland, FR = France). Bold italic values correspond to the number of species.

are volunteers (82.3% +/- 5.8), investing on average 1939.6 +/- 534.6 days per year, visiting 1098.7 +/- 357.1 sites (median = 560) on average 7.3+/-3.7 times (median = 2; see Figure 3 for more details).

The main factors of ecological change that coordinators considered that they could assess with their monitoring data were land use change in small and medium programs (Table 1). In large programs, a majority of programs monitored land use changes and climate change impacts (Table 1). The distributions of the ecological factors monitored differed significantly between the differently sized monitoring programs (χ^2_2 = 6.879, p = 0.032; Table 1). In all three categories of monitoring programs, population trends were the first target of the monitoring. Community trends were least monitored across all programs sizes (Table 1). Most of the small and medium programs were scientific programs. In medium programs, many also had a management motivation (34.6%), while the large programs included 34.4% scientific, 28.1% political and 25% management programs (χ^2_2 = 1.294; p = 0.523; Table 1).

In small programs, sites were mainly chosen through expert knowledge (Table 1). In medium programs, sampling was most frequently exhaustive or based on site choice according to expert knowledge (Table 1). In large programs, random sampling and site choice by expert knowledge was most frequently employed (Table 1). Whether monitoring was conducted within and/or outside a protected area was independent of monitoring program size, as was the field data type that were most frequently used (Table 1). The issue of detection probability was neglected in all types of programs (χ^2_2 = 0.092; p = 0.955), ranging from 35% of the total number of small programs to 46.7% (large programs). The same result was found for the application of stratified sampling (χ^2_2 = 2.656; p = 0.265; Table 1). In small and medium programs, basic statistics (descriptive statistics or correlations) were most frequently used, while large programs may have more frequently used more advanced statistics (χ^2_2 = 3.348; p = 0.188).

Several monitoring programs focused on one of the three bird species groups, raptors (N = 16), waterbirds (N = 44), and songbirds and near passerines only (N = 27). Raptor programs monitored 4.8 +/- 1.3 species (median = 3) with 42.6 +/- 16.3 persons (median = 11), which were mainly professionals (58% +/- 11.6; median 14.3%), investing 392.5 +/- 127.9 days per year, visiting 164.1 +/- 91.5 sites (median = 23) on 3.7 +/- 1.0 times (median = 2). The majority of raptor programs did not consider detection probability nor stratification (87.5%), but followed an exhaustive sampling design (56%) and analyzed data with basic statistics only (12.5%). Waterbird programs monitored 16.9 +/- 4.1 species (median = 1) with 138.9 +/- 52.6 persons (median = 30), which were mainly volunteers (52% +/- 6.2; median 60.0%), investing on average 869.4 +/- 374.5 days per year, visiting 294.3 +/- 111.1 sites (median = 50) on average 22.3 +/- 12.1 times (median = 2). In waterbird programs 25% considered detection probability and 20.5% stratified their sample. The sampling design was either exhaustive (36.4%) or following expert knowledge (47.7%), while only 9% employed random or systematic sampling. Songbirds and near passerine programs monitored 26.0 +/- 8.0 species (median = 12) with 400.7 +/- 369.3 persons (median = 22), which were mainly volunteers (59.8% +/- 8.4; median 87%), investing on average 222.2 +/- 72.5 days per year, visiting 409.3 +/-

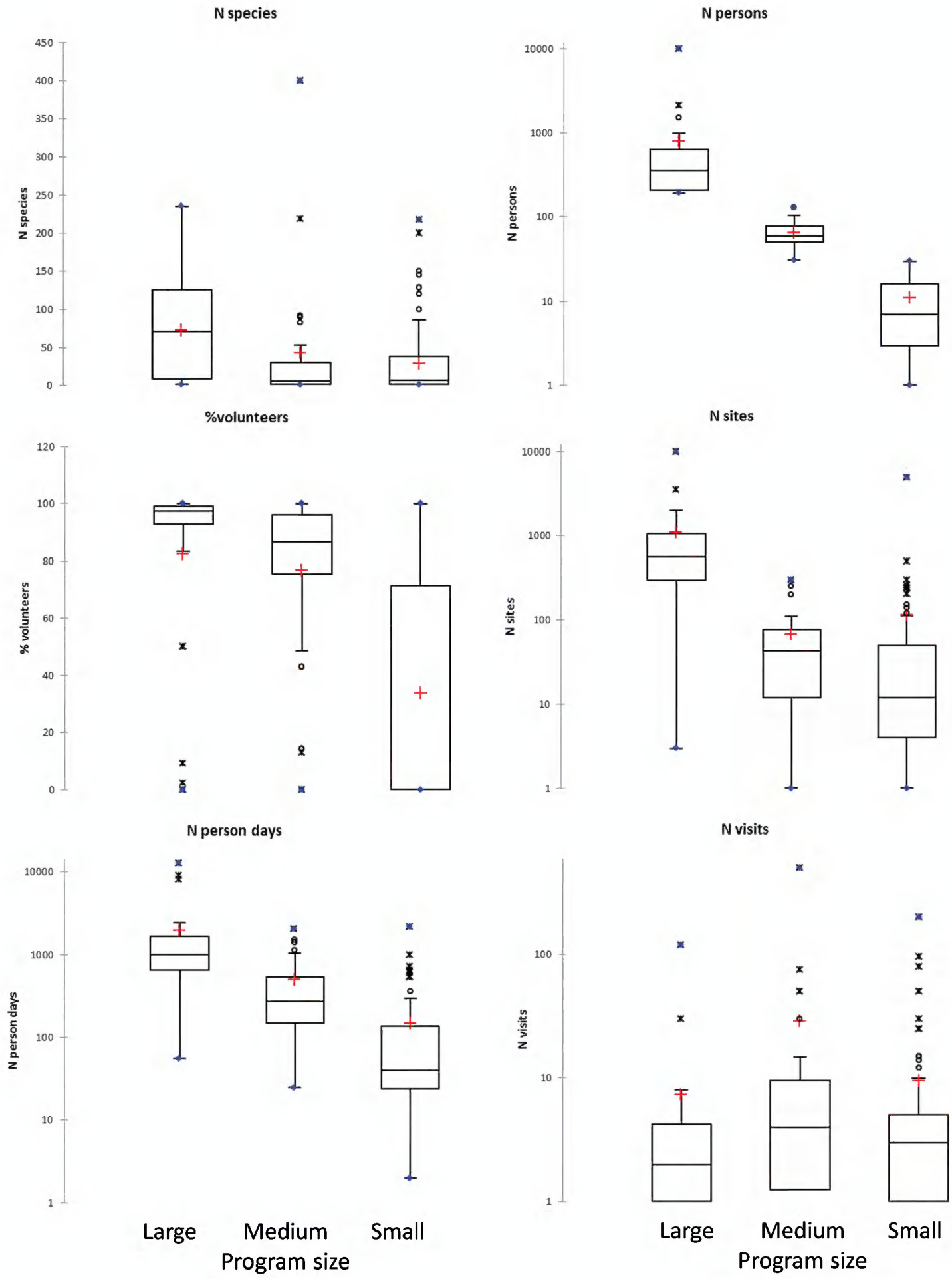

Figure 3. Univariate boxplots on the sampling effort and proportion of volunteers for small, medium and large European bird-monitoring programs (the size of a monitoring scheme was defined mainly by the number of people involved).

Table 1. Summary of responses of European bird species programs to questions on the motivation and aims, and the sampling design. The values are given for small, medium and large programs as defined by the number of people involved in the monitoring (see also text). For more information on the questions, see the supplementary material ESM 1, which is available online.

		Small N (%)	Medium N (%)	Large N (%)
Motivation and aim				
	causes of change a program monitors			
	Land use	46 (56.8%)	15 (57.7%)	20 (62.5%)
	Climate change	20 (24.7%)	11 (42.3%)	17 (53.1%)
	Habitat fragmentation	19 (23.5%)	3 (11.5%)	8 (25.0%)
	Pollution	18 (22.2%)	3 (11.5%)	9 (28.1%)
	Invasive species	9 (11.1%)	4 (15.4%)	8 (25.0%)
	Reason for launching a program			
	Scientific	33 (40.7%)	11 (42.3%)	11 (34.4%)
	Political	24 (29.6%)	4 (15.4%)	9 (28.1%)
	Management or restoration	21 (25.9%)	9 (34.6%)	8 (25.0%)
	Other	2 (2.5%)	2 (7.7%)	4 (12.5%)
	Monitored trends			
	Population trends	79 (97.5%)	26 (100%)	31 (96.9%)
	Distribution trends	44 (54.3%)	16 (61.5%)	25 (78.1%)
	Community trends	41 (50.6%)	5 (19.2%)	7 (21.9%)
Sampling design				
	Choice of site			
	Personal or expert knowledge	46 (65.0%)	9 (37.5%)	12 (37.5%)
	Exhaustive sampling	26 (32.5%)	10 (41.7%)	4 (12.5%)
	Systematic sampling	4 (5.0%)	2 (8.3%)	2 (6.3%)
	Other	3 (3.7%)	3 (12.5%)	5 (15.6%)
	Random sampling	1 (1.3%)	-	9 (28.1%)
	Field data type			
	Age	-	-	1 (3.1%)
	Counts	53 (65.4%)	14 (53.8%)	24 (75.0%)
	Mark-recapture	16 (19.7%)	7 (26.9%)	2 (6.3%)
	Presence-absence	8 (9.9%)	3 (11.5%)	2 (6.3%)
	Phenology	4 (4.9%)	2 (7.7%)	3 (9.4%)
	Monitored area legally protected			
	Both, within and outside a reserve site	49 (61.3%)	17 (65.4%)	28 (87.5%)
	Within reserve site	21 (26.3%)	3 (11.5%)	-
	Outside reserve site	10 (12.5%)	6 (23.1%)	4 (12.5%)
Data processing				
	Detection probability	28 (35.0%)	9 (36.0%)	14 (46.7%)
	Stratification	18 (22.2%)	8 (30.7%)	12 (37.5%)

383.7 sites (median = 10) on average 11.8 +/- 3.4 times (median = 2). Songbirds and near passerine programs accounted for detection probability in 63% and used stratification in 26% of the programs. Most of the programs (74.1%) used a sampling design following expert knowledge and used advanced statistics (63%) for data analysis.

Sampling and data processing

The field data type largely depended on the program objective ($\chi^2_2 = 10.11$; $p = 0.006$): programs with a scientific motivation more frequently employed mark-recapture studies (35%) as compared to management/restoration programs (10.5%), while politically motivated programs did not employ mark-recapture methods at all. Conversely, counts were used less frequently in scientific programs (46%) as compared to management/restoration programs (71%) and programs with a political interest (84%).

Site choice methodology was related to the proportion of volunteers involved ($\chi^2_1 = 4.67$; $p = 0.031$). Programs with more professionals than volunteers employed systematic sampling or chose sites based on expert knowledge, while programs with exhaustive or random sampling were dominated by volunteers. Consideration of detection probability was related to the program objective ($\chi^2_2 = 16.71$; $p < 0.001$): scientifically oriented programs accounted more often for detectability than other programs, although still 46% of the scientifically motivated programs ignored the problem of detection probability as did 66.7% for management programs and 82.8% for political programs. Stratification was used in few programs (31% of scientific programs; 23.7% of management programs; 16.2% of politically motivated programs; $\chi^2_2 = 2.043$; $p = 0.36$).

Advanced statistics (i.e. GLM, or Generalized Additive Models) were more likely used for data analysis with increasing total sampling effort (number of person days) and varied with the program objective (respectively, $\chi^2_1 = 11.58$; $p < 0.001$; $\chi^2_2 = 14.76$; $p < 0.001$); 62.5% of the scientific programs used advanced statistics, 47% in management programs, and 23.5% in politically motivated programs. The level of statistical data processing (use of basic or advanced statistics) was not related to the sampling design ($\chi^2_4 = 6.04$; $p = 0.196$).

Discussion

The majority of programs of our database comprised of small programs, i.e. monitoring few bird species with few people. These programs were homogeneous in terms of practices for monitoring bird populations on a local scale using counts or even capture-mark-recapture data to monitor population trends in detail. Capture-mark-recapture data were usually collected in scientifically motivated programs at sites chosen by experts. Fewer programs were medium-sized, focusing on populations on a local to regional scale, using count data and an exhaustive sampling design. The large monitoring programs sampled count data, while selecting sites either randomly or following expert opinion.

Monitoring programs share the common desire to determine what changes are occurring in bird populations and why these changes occur. Programs at different scales are needed to address these questions, although their primary aims may differ depending on the scale of implementation. Large-scale monitoring programs across biogeographic regions, countries or a continent are usually designed to determine if population changes are occurring. However, the design of large-scale programs is too coarse to provide information on changes at specific sites or to provide direct information on the causes of population change. Here, small-scale monitoring programs are needed to analyze why population change is occurring at specific sites. Such local-scale data can then feed into management and conservation actions for specific sites. With these differences in mind, it is little surprising that population trends are by far the most frequently monitored trend, regardless of the size of the monitoring program.

Due to the aims of a local scientific program, few employed random sampling, while site selection was done according to expert knowledge. While such a design is suitable for specific (scientific) questions, a subjective sampling effort in general must be considered as a poor design for a monitoring program since it provides a biased coverage of the mechanisms at play, without characterizing the biases. Surprisingly, our data suggest that random sampling, while highly recommended, was employed by only 28% of the large-scale programs and hence 72% did not follow the recommendations of good monitoring practices (Gaines et al. 1999; Yoccoz et al. 2001; Nichols and Williams 2006). Note that the national Breeding Bird Surveys (Gibbons 2000) usually employ randomized or semi-randomized sampling but large-scale, national bird-monitoring programs formed a minority of the schemes in our database.

Concerning data collection, bird-monitoring data were usually counts, largely dominating across all monitoring programs in our database. Resource intensive capture-mark-recapture studies (Voříšek et al., 2008) were usually conducted at local and regional scales. The small and locally focused monitoring programs, however, need to be put into a large-scale perspective to determine if changes are due to local or external factors. Such a consideration is important for a generalization of trends across geographic and temporal scales. Therefore, it is important that the results from small monitoring programs are interpreted relative to changes at the population level. They can then serve as benchmark sites for large-scale monitoring programs, thereby providing in-depth information at specific sites (Downes et al. 2005; Henry et al. 2008). Our analysis shows that the potential of such an integration of local and small monitoring programs on a larger-scale is high, given that the homogeneity of the different parameters analyzed in our sample of small monitoring programs was comparably high. Integrating the monitoring data of the 81 small monitoring programs could yield a remarkably good coverage and profound insight of local impacts on bird populations across Europe.

In respect to the determination of the causes of change in population trends, it is also important to monitor sites in and outside of protected areas since the pressures are different. Our data suggests that this notion is well implemented in bird-monitoring in Europe, improving our ability to generalize results by comparing population changes within and outside of protected areas. Such comparisons are of special importance to

disentangle large-scale factors (such as climate change), from more local effects (such as habitat fragmentation and pollution for instance).

Concerning sampling stratification, we also found a difference between the differently sized programs, which is likely to be related to the differences in the aim and design of small to large programs. In small programs, stratified sampling was applied in only 22.2%, while in large programs the proportion raised to 37.5% (30.7% in medium programs). For local and regional programs such a proportion might be sufficient since homogeneity of the sample population is higher at a smaller scale. In contrast, stratified sampling must be employed more frequently in large-scale programs due to limited resource and sampling disequilibrium between potential strata.

The largest deficit in the consideration of recommendations was the lack of repeated sampling to account for detection probability. Only little more than a third of programs employed repeated sampling, usually programs with a scientific motivation. Programs with management objectives and with a political motivation employed repeated sampling even less often, making them more prone to misinterpretation of trends that may be due to variations in detection probability (Pellet and Schmidt 2005; Schmidt 2005; Kery et al. 2006). Here, large programs performed the best. In programs on a local scale, detection probability might not be considered due to two reasons: (i) detection of a focal species is considered sufficient as sites are visited more frequently or enough so that trends are not biased (but see Archaux et al. 2011), or (ii) the statistical analysis needed to model detection probability appears too complex. This would be coherent with the fact that small and medium programs usually employed only basic statistics, while large programs used advanced statistics to value their large datasets. Hence, recommendations of good monitoring practice are only followed by a minority of the programs, with many consequences for the interpretation of data, especially in politically motivated programs.

Generally, our data show that there is a huge variety of monitoring practices across all monitoring programs, among and within bird species groups, partly explained by the program objective, and the scale of the implementation of programs. It appears to be justified to recommend that bird-monitoring in Europe may step up the effort in methodological implementation of monitoring recommendations (Voříšek et al. 2008) to produce more standardized bird-monitoring data. Such an effort would increase the potential uses of these data, and particularly the potential for the integration of data at large geographical scales (Downes et al. 2005; Henry et al. 2008; Pereira et al. 2010).

Resource limitations and volunteer-based monitoring

The culture of bird-monitoring was born and propagated by visionary bird watching and naturalist amateurs, led by skilled professionals. This enabled the founding of long-term databases with minimal funding. Due to this historical contingency, the involvement of volunteers in monitoring is still key to maximize the sampling effort and to acquire a large-scale image of changes in bird diversity (Engel and Voshell Jr. 2002;

Bell et al. 2008; Schmeller et al. 2009). At first sight, our survey concurs with the common belief that an optimized sampling design is poorly compatible with massive volunteer involvement. The recommended stratified and/or random spatial sampling (Yoccoz et al. 2001; Vořišek et al. 2008) is used in only a rather small proportion of the monitoring programs, suggesting that coordinators believe that they cannot impose sophisticated sampling designs if they want to attract large numbers of volunteers. However, our survey shows that 14% of programs have successfully used random sampling designs, further improved by the use of advanced statistical analysis, showing that volunteer involvement is actually compatible with good monitoring practices (Schmeller et al. 2009). This optimization of monitoring constraints is well illustrated by Vořišek and colleagues (2003; 2008), in their overview of the national common bird surveys in Europe (http://www.ebcc.info/pecbm.html). Random sampling could be achieved in most of the programs once volunteers see the advantages of random *versus* opportunistic sampling (Buckland et al. 2005). We believe that the key to improving average monitoring practices is the involvement of skilled biologists, engaging in training and effective communication regarding sampling design and data processing and analysis.

Recommendations

In the monitoring literature a three-phase approach is described for the process of biodiversity monitoring, (i) identifying monitoring questions and aims, (ii) identifying the most suitable monitoring methods, and (iii) interpreting monitoring data (Gaines et al. 1999; Yoccoz et al. 2001; Vořišek et al. 2008). For most monitoring programs, the best data type to be collected is count data, which enable management actions and secure an early warning for conservation and policy. More sophisticated methods like capture-mark-recapture studies can then be employed to explore more specific scientific questions (Lepetz et al. 2009). Count data has the best trade-off between resource use for data collection and the quantity of information contained in the data. Further, monitoring could be stratified to optimize resource allocation between independent samples (i.e., sites), and employ random (or systematic) sampling to secure an unbiased spatial coverage. Importantly, detection probability needs to be accounted for since even low differences in detection probability between site or years can induce spurious conclusions (Archaux et al. 2011). It means that repetitive sampling of the same sites within a year should be the rule. In case of limited manpower, Vořišek and colleagues (2008), among others, recommended to maximize the number of samples, even at the expense of the size of each sample, so that the precision of population estimates remains the highest possible, allowing a better coverage of the different sources of heterogeneity which in turn can also limit the bias. For the statistical analysis, it is advantageous to not only use descriptive statistics or simple correlation analysis, as these techniques do not optimize the extraction of the information contained in the monitoring data. There is a range of different free software packages available, which could be used to do advanced statistics with count and capture-mark-recapture data (e.g. TRIM, MARK, and several R packages). For further valoriza-

tion of monitoring data, coordinators need to consider data integration across different monitoring programs. Therefore, guidelines for data integration across programs need to be clear, comprehensible and accessible to monitoring coordinators (Henry et al., 2008). Further, more collaborations between monitoring programs at different scales need to be established, so that the numerous datasets currently not included in the evaluation of trends in bird populations might be better considered in the future. Finally, monitoring coordinators may wonder how to attract volunteer monitors for a specific program to increase the manpower without over-stretching the financial budget. Several factors define a successful volunteer involvement (Bell et al. 2008; Schmeller et al. 2009; Vandzinskaite et al. 2010): (i) the socio-political background influences levels of participation, (ii) different recruitment strategies are needed for retention of volunteers, (iii) keep volunteers informed, (iv) carefully consider relationships between professionals and volunteers, and (v) collaborate with other monitoring programs to add value.

Acknowledgements

This contribution was supported by funds of the EuMon STREP EU-project (EU-Commission contract number 6463). We would like to thank Ian Burfield for his substantial contribution to this publication and language revisions, the editor and two anonymous reviewers for their helpful comments. We further thank all the contributors of monitoring program descriptions, who spent their precious time answering our questions, and our colleagues who helped in assembling this database, particularly P. Nowicki, E. Budrys, M. Kotarac, B. Gruber, and B. Ivanek. This database is publicly available at http://eumon.ckff.si/monitoring/search.php.

References

Archaux F, Henry PY, Gimenez O (2011) When can we ignore the problem of imperfect detection in comparative studies? Methods in Ecology and Evolution 3: 188–194. doi: 10.1111/j.2041-210X.2011.00142.x

Barclay S, Todd C, Finlay I, Grande G, Wyatt P (2002) Not another questionnaire! Maximizing the response rate, predicting non-response and assessing non-response bias in postal questionnaire studies of GPs. Fam Pract 19: 105–111. doi: 10.1093/fampra/19.1.105

Bell S, Marzano M, Cent J, Kobierska H, Podjed D, Vandzinskaite D, Reinert H, Armaitiene A, Grodzińska-Jurczak M, Muršič R (2008) What counts? Volunteers and their organisations in the recording and monitoring of biodiversity. Biodiversity and Conservation 17: 3443–3454. doi: 10.1007/s10531-008-9357-9

Buckland ST, Magurran AE, Green RE, Fewster RM (2005) Monitoring change in biodiversity through composite indices. Philosophical Transactions of the Royal Society of London Series B Biological Sciences 360: 243–254. doi: 10.1098/rstb.2004.1589

Burfield I, van Bommel F, Gallo-Orsi U, Nagy S, Orhun C, Pople R, van Zoest R, Callaghan D, International. B (2004) Birds in Europe: Population Estimates, Trends and Conservation Status. BirdLife International, Cambridge.

Danielsen F, Balete DS, Poulsen MK, Enghoff M, Nozawa CM, Jensen AE (2000) A simple system for monitoring biodiversity in protected areas of a developing country. Biodiversity and Conservation 9: 1671–1705.

Donald PF, Sanderson FJ, Burfield IJ, Bierman SM, Gregory RD, Waliczky Z (2007) International Conservation Policy Delivers Benefits for Birds in Europe. Science 317: 810–813. doi: 10.1126/science.1146002

Downes C, Bart J, Collins B, Craig B, Dale B, Dunn E, Francis C, Woodley S, Zorn P (2005) Small-Scale Monitoring: Can It be Integrated with Large-scale Programs? In: Ralph CJ, Rich TD (Eds) Bird Conservation Implementation and Integration in the Americas: Proceedings of the Third International Partners in Flight Conference March 20–24, 2002. Pacific Southwest Research Station, Asilomar, California, 993–996.

Engel SR, Voshell Jr JR (2002) Volunteer Biological Monitoring: Can It Accurately Assess the Ecological Condition of Streams? American Entomologist 48: 164–177.

Gaines WL, Harrod RJ, Lehmkuhl JF (1999) Monitoring biodiversity: quantification and interpretation. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland.

Gibbons D (2000) Development of Pan-European breeding bird-monitoring. Ring 22: 25–33.

Gregory RD, Strien A, Vorisek P (2006) Using birds as indicators of environmental change in Europe. Journal of Ornithology 147: 16 pp.

Gregory RD, van Strien A, Vorisek P, Meyling AWG, Noble DG, Foppen RPB, Gibbons DW (2005) Developing indicators for European birds. Philosophical Transactions of the Royal Society B-Biological Sciences 360: 269–288. doi: 10.1098/rstb.2004.1602

Henry PY, Lengyel S, Nowicki P, Julliard R, Clobert J, Celik T, Gruber B, Schmeller DS, Babij V, Henle K (2008) Integrating ongoing biodiversity monitoring: potential benefits and methods. Biodiversity and Conservation 17: 3357–3382. doi: 10.1007/s10531-008-9417-1

Kery M, Schmid H (2004) Monitoring programs need to take into account imperfect species detectability. Basic and Applied Ecology 5: 65–73. doi: 10.1078/1439-1791-00194

Kery M, Spillmann JH, Truong C, Holderegger R (2006) How biased are estimates of extinction probability in revisitation studies? Journal of Ecology 94: 980–986. doi: 10.1111/j.1365-2745.2006.01151.x

Klvanová A, Voríšek P (2007) Review on large-scale generic population monitoring programs in Europe 2007. Bird Census News 20: 50–56.

Kull T, Sammul M, Kull K, Lanno K, Tali K, Gruber B, Schmeller D, Henle K (2008) Necessity and reality of monitoring threatened European vascular plants. Biodiversity and Conservation 17: 3383–3402. doi: 10.1007/s10531-008-9432-2

Lengyel S, Déri E, Varga Z, Horváth R, Tóthmérész B, Henry P-Y, Kobler A, Kutnar L, Babij V, Seliškar A, Chysoula C, Papastergiadou E, Gruber B, Henle K (2008) Habitat monitoring in Europe: a description of current practices. Biodiversity and Conservation 17: 3327–3339. doi: 10.1007/s10531-008-9395-3

Lepetz V, Massot M, Schmeller DS, Clobert J (2009) Biodiversity monitoring: some proposals to adequately study species' responses to climate change. Biodiversity and Conservation 18: 3185–3203. doi: 10.1007/s10531-009-9636-0

Lindenmayer D, Likens G (2010) The science and application of ecological monitoring. Biological Conservation 143: 1317–1328. doi: 10.1016/j.biocon.2010.02.013

Lindenmayer DB, Likens GE (2009) Adaptive monitoring: a new paradigm for long-term research and monitoring. Trends in Ecology & Evolution 24: 482–486. doi: 10.1016/j.tree.2009.03.005

Male TD, Bean MJ (2005) Measuring progress in US endangered species conservation. Ecology Letters 8: 986–992. doi: 10.1111/j.1461-0248.2005.00806.x

Marsh DM, Trenham PC (2008) Current Trends in Plant and Animal Population Monitoring. Conservation Biology 22: 647–655. doi: 10.1111/j.1523-1739.2008.00927.x

Nichols JD, Williams BK (2006) Monitoring for conservation. Trends in Ecology & Evolution 21: 668–673. doi: 10.1016/j.tree.2006.08.007

Pellet J, Schmidt BR (2005) Monitoring distributions using call surveys: estimating site occupancy, detection probabilities and inferring absence. Biological Conservation 123: 27–35. doi: 10.1016/j.biocon.2004.10.005

Pereira HM, Belnap J, Brummitt N, Collen B, Ding H, Gonzalez-Espinosa M, Gregory RD, Honrado J, Jongman RH, Julliard R, McRae L, Proenca V, Rodrigues P, Opige M, Rodriguez JP, Schmeller DS, van Swaay C, Vieira C (2010) Global biodiversity monitoring. Frontiers in Ecology and the Environment 8: 459–460. doi: 10.1890/10.WB.23

Schmeller DS, Henry PY, Julliard R, Gruber B, Clobert J, Dziock F, Lengyel S, Nowicki P, Deri E, Budrys E, Kull T, Tali K, Bauch B, Settele J, Van Swaay C, Kobler A, Babij V, Papastergiadou E, Henle K (2009) Advantages of Volunteer-Based Biodiversity Monitoring in Europe. Conservation Biology 23: 307–316. doi: 10.1111/j.1523-1739.2008.01125.x

Schmidt BR (2005) Monitoring the distribution of pond-breeding amphibians when species are detected imperfectly. Aquatic Conservation-Marine and Freshwater Ecosystems 15: 681–692. doi: 10.1002/aqc.740

Taylor MFJ, Suckling KF, Rachlinski JJ (2005) The Effectiveness of the Endangered Species Act: A Quantitative Analysis. BioScience 55: 360–367. doi: 10.1641/0006-3568(2005)055[0360:TEOTES]2.0.CO;2

Tucker GM, Heath MF (1994) Birds in Europe: Their conservation status. Birdlife International, Cambridge, UK, 600 pp.

Vandzinskaite D, Kobierska H, Schmeller DS, Grodzińska-Jurczak M (2010) Cultural Diversity Issues in Biodiversity Monitoring—Cases of Lithuania, Poland and Denmark. Diversity 2: 1130–1145. doi: 10.3390/d2091130

Voříšek P, Klvaňova A, Wotton S, Gregory RD (2008) A best practice guide for wild bird-monitoring programs.

Voříšek P, Marchant J (2003) Review of large-scale generic population monitoring programs in Europe. Bird Census News 16: 16–40.

Yoccoz NG, Nichols JD, Boulinier T (2001) Monitoring of biological diversity in space and time. Trends in Ecology & Evolution 16: 446–453. doi: 10.1016/S0169-5347(01)02205-4

Appendix

Supplementary material to article European Bird-monitoring - an overview. (doi: 10.3897/natureconservation.2.3644.app) File format: MS Word Document (doc).

Explanation note: The questionnaire was designed to assess how biodiversity monitoring schemes were carried out and what the motivation was to launch that scheme.

Note: The following is a transcript of the questions coordinators answered in the online questionnaire available at: http://eumon.ckff.si/monitoring. The online version also contains clarifications and explanatory notes.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Citation: Schmeller DS, Henle K, Loyau A, Besnard A, Henry P-Y (2012) Bird-monitoring in Europe – a first overview of practices, motivations and aims. Nature Conservation 2: 41–57. doi: 10.3897/natureconservation.2.3644.app